GRESHAM QUANT
We Know Systematic®

Trend: You Lead, I'll Follow

There are many ways to trend-follow an asset, but most of them boil down to taking risk which is proportional to the weighed sum of recent (vol adjusted) historical returns.

$$\text{sig}(t) = \sum_{i=1}^{N} w_i r_{t-i}$$

It's so simple it shouldn't work. Doesn't every investment disclaimer remind us that past performance is not indicative of future returns? And yet it does. There are both economic reasons (price discovery takes time) and behavioural reasons (everyone loves to buy winners) why, but as mathematicians we distil all these complicated reasons to the following:

- Trend will make money if, for whatever reason, the underlying returns have a persistent drift
- Trend will make money if, for whatever reason, past returns really do correlate to future returns

We have looked at the impact of long-term drift on trend's performance in our "How good is trend following at following trend" blog, so today let's look at how the positive autocorrelation of returns impact trend's performance.

Risks and rewards

In order to evaluate our signals performance, we need to know two things:

- how much money do we expect it to earn?
- how much risk does the signal take?

The ratio between expectation and risk is the daily "signal to noise" ratio, and that scales to an annual Sharpe by a factor of 16 (if there are 256 business days in a year).

If daily returns are N(μ,1) and corr(r_t, r_{t-i}) = ρ_i, taking the expectation is easy:

$$\mathbb{E}(\text{sig}(t)r(t)) = \sum_{i=1}^{N} w_i \mathbb{E}(r(t)r(t-i)) = \sum_{i=1}^{N} w_i(\mu^2 + \rho_i)$$

As for risk, we usually measure risk as the root mean squared (RMS) of the signal, i.e.

$$\sqrt{E(\text{sig}(t)^2)} = \sqrt{\mathbb{E}\left[\sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j r(t-i) r(t-j)\right]}$$

We can break the summation into i=j and i≠j

$$\mathbb{E}\left[r(t-i)r(t-i)\right] = \mu^2 + 1$$

$$\mathbb{E}\left[r(t-i)r(t-j)\right]\Big|_{i\neq j} = \mathbb{E}(r(t-i))\mathbb{E}(r(t-j)) + \rho(|i-j|) = \mu^2 + \rho(|i-j|)$$

So

$$\mathbb{E}(\text{sig}^2) = \sum_{i=1}^{N} w_i^2 + 2\sum_{i=1}^{N}\sum_{j=1}^{i-1} w_i w_j \rho(|i-j|) + \left(\sum_{i=1}^{N} w_i \mu\right)^2$$

All three of these terms are interpretable:

- The first two are the covariance of the signal, the first being the variance of each of the individual returns and the second the additional covariance due to autocorrelation
- The last term is just $E(\text{sig})^2$ and is the contribution we get from the fact that the signal has a non-zero expected value.

How do trend followers manage risk?

Trend followers will try and ensure that $E(\text{sig}^2) = 1$. This means that risk we assign to different trend predictors are comparable in size.

If we were omniscient and knew the drift or the autocorrelation, we would choose our weights precisely. Of course, if we knew the drift and autocorrelation in advance, we would be very rich indeed. But we don't. And we're not. So instead, we ignore the last two terms, and will choose weights such that

$$\sum_{i}^{N} w_i^2 = 1$$

As it happens, in real life, the last two terms in the above equation are small, so trend followers do a very decent job at sizing their bets. In fact, ignoring the last two terms works in favour of trend followers: we tend to over-allocate risk (by having larger than average signals) precisely to markets that have drift and positive autocorrelation.

Less Equations, More Explanations

Let's look at a simple example to demonstrate the general case. We will look at a simple three-month (64-day) rolling average signal (all weights are 1/8 so that $\sum w_i^2 = 1$), and we will assume a constant autocorrelation for any lag 1…64 and zero autocorrelation for longer lags.

So how much risk is the signal taking?

$$\mathbb{E}(\text{sig}^2) = 1 + 64\mu^2 + 63\rho$$

And what about expected profit?

$$\mathbb{E}(\text{sig} * r(t)) = 8(\mu^2 + \rho)$$

Then the daily signal-to-noise ratio is (approximately)

$$\frac{8(\mu^2 + \rho)}{\sqrt{1 + 64\mu^2 + 63\rho}}$$

Reality check

In real life, the average daily auto correlation, ρ, is small -- perhaps 0.2% -- while a reasonably strong trend would equate to a yearly drift of about 1-standard deviation of annual volatility, so daily drift μ is about 1/16 of daily volatility.

A three-month-rolling-average trend follower would then expect to make (per day):

$8 * (1/16)^2 + 8 * 0.2\% = 0.03125 + 0.016 = 0.04725$

And its risk would be $\sqrt{(1 + 0.25 + 0.126)} \approx 1.17$

These are small daily expected gains but scaled over a year, we multiply everything by 16, and these suddenly become a Sharpe 0.6 system, with about 0.4 Sharpe coming from drift and a Sharpe 0.2 coming from autocorrelation.

Now we can draw some conclusions. Let us start with risk: There is a correction term for the correlations in the signal's risk, but unless autocorrelation average around 1%, the contribution to risk is not that meaningful. Similarly, only for large trends the drift term starts to contribute meaningfully to the risk we end up taking, reducing our overall realised Sharpe.

More interesting though is what drives expected gains. Autocorrelation impacts gains in a meaningful way, indeed, for the (quite realistic) example above, expected returns due to autocorrelation are half the size of returns we associate with a decent drift.

So autocorrelation is crucial when it comes to understanding trend's expected performance.

Total autocorrelation (TAC)

Total autocorrelation, autocorrelation summed across all lags, is how we think about **trend quality**. We talk a lot about how **much** markets trend, but the above gives us the language to talk about how **well** markets trend.

Annualizing the above equation, we notice that 63ρ is pretty much the total autocorrelation (TAC), and daily drift (as a fraction of daily vol) is 1/16th of the annual drift (as a fraction of annual volatility) so:

$$\text{Daily Signal-to-Noise Ratio} \approx \frac{8(\mu_{\text{annual}}^2/256 + \text{TAC}/64)}{\sqrt{1+\mu_{\text{annual}}^2/4 + \text{TAC}}}$$

$$\text{Annual Sharpe Ratio} \approx \frac{\mu_{\text{annual}}^2/2 + 2 * \text{TAC}}{\sqrt{1+\mu_{\text{annual}}^2/4 + \text{TAC}}} \qquad \text{(HGIT)}$$

HGIT stands for How-Good-Is-Trend, and the equation does exactly[1] what it says on the tin. If you want to make money in trend-following you need some combination of both quantity (μ) and quality (TAC). If you have enough extrinsic trend, you can get away with lower autocorrelation but in the absence of trends, it's positive autocorrelation that will save you.

The Proof is in the Pudding

Let's look at some Monte-Carlo simulations, simulating *annual* drift (as a fraction of annual volatility).

Here we are plotting both the theoretical HGIT approximation of trend performance and a Monte Carlo simulation of trend-following on a single market. The underlying daily market returns are generated from N(μ,1) where we sample a range for μ to build up a picture of how this looks for different drifts. We also introduce autocorrelation of TAC/64 on the most recent 64 days. Each simulation will give rise to an observed drift over a year, and we will plot the simulation's Sharpe against the underlying asset's realised Sharpe, which comes from the drift of that asset over the course of the simulation.

Let's look first at the case with zero autocorrelation, i.e. daily returns are drawn from IID N(μ,1). Our maths (once annualized) predicts a 3m trend follower:

$$\text{Sharpe}_{3m}(\mu) = \frac{\mu^2/2}{\sqrt{1+\mu^2/4}}$$

[1] HGIT is not exact, E(return/risk) is most certainly not E(return)/√E(risk²). Crucially, we ignored the fact that risk and returns are positively correlated, decreasing the overall Sharpe expectation slightly.

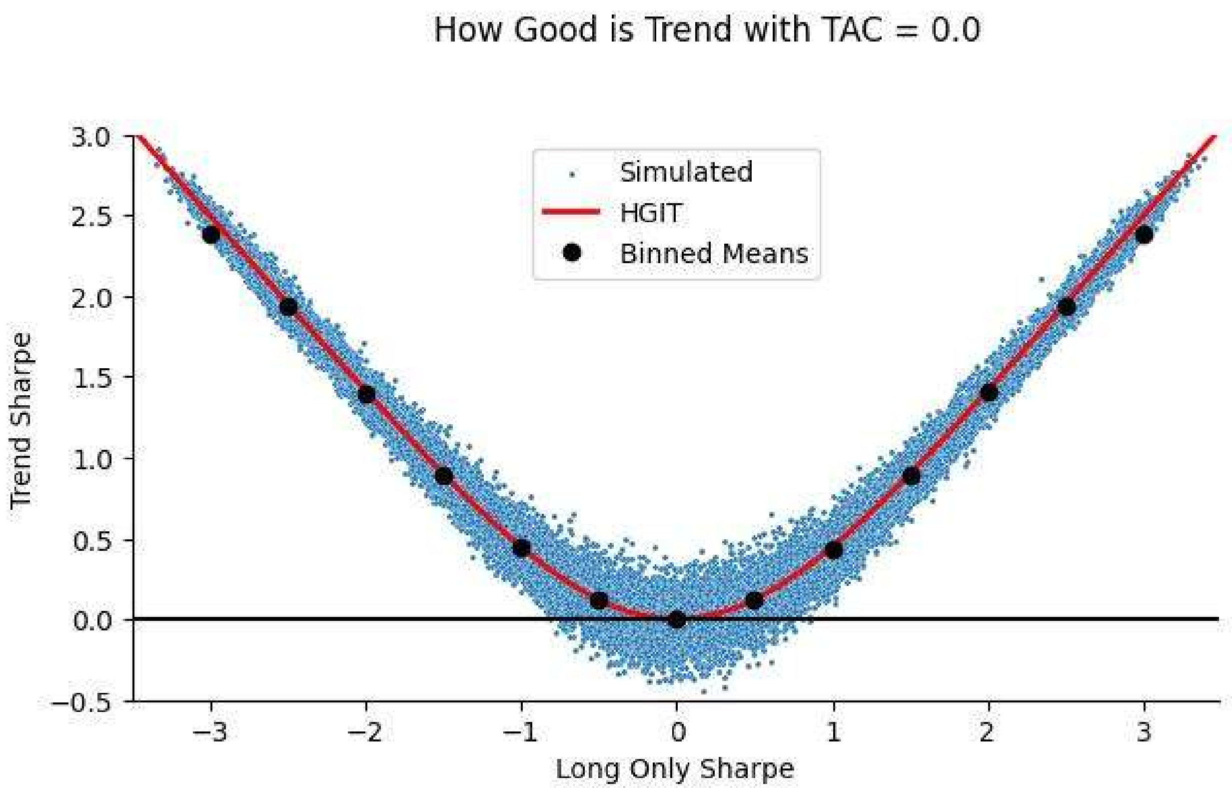

Figure 1: Realized Sharpe simulations: unrealistically, expected Sharpe is always positive! This is of nonsense and in real life there is a "haircut" we have to pay due to trading costs, non-constant drifts and life rarely fitting toy models

And thankfully, the Monte-Carlo simulations confirm our intuition.

Now let's look at the case where we have some positive autocorrelation in the returns. Let's take a total autocorrelation of 0.1, evenly distributed over the first 64 lags. HGIT predicts:

$$\text{Sharpe}_{3m}(\mu, \text{TAC} = 0.1) = \frac{\mu^2/2 + 0.2}{\sqrt{1 + \mu^2/4 + 0.1}}$$

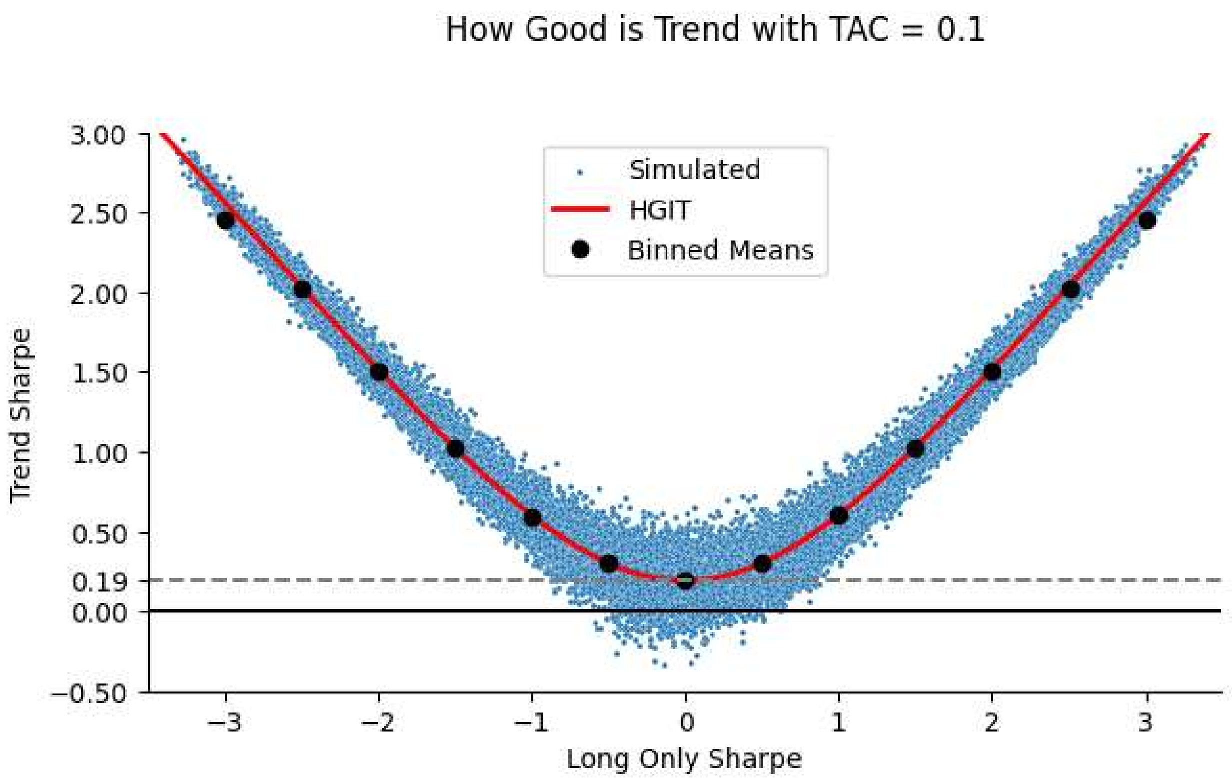

Figure 2: Monte-Carlo simulations of realized Sharpes for a 3m trend, when total autocorrelation is at 0.1

Lo and behold, the simulations agree with our formulae! Crucially, even for *zero* realised drift in the underlying, the average simulated Sharpe is higher by 0.19.

With Great Autocorrelation Comes Great Variability

You could be forgiven for not noticing this on the plots above, but something else is happening, besides the shift upwards. Even though the daily returns were sampled with unit variance, the standard deviation of the annual long-only Sharpe is a bit higher. When autocorrelation is positive, for a given daily volatility, we get a higher yearly volatility as trend reinforces itself! This is excellent news for trend followers: not only do we make more money from a given drift, we also spend more time in the tails!

We can calculate this increased variance easily. So easily that we leave it as an exercise! It turns out that annual volatility increases with TAC approximately like $\sqrt{(1+2\ TAC)}$

We run our Monte Carlo and look at the standard deviation of realised observed drifts as we increase the total autocorrelation. The change in the distribution of μ agrees with our intuition.

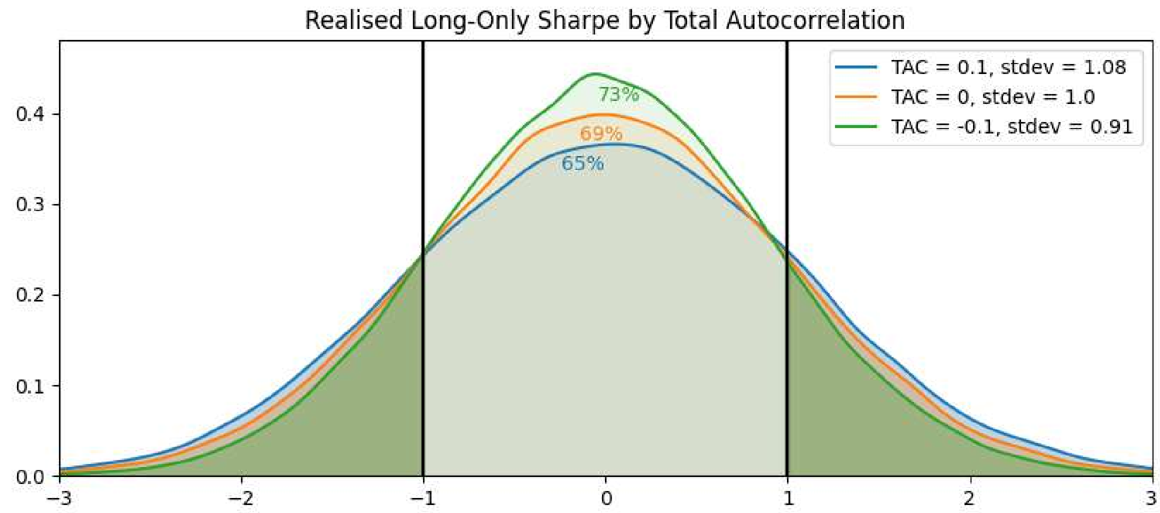

Figure 3: Realised drifts by autocorrelation. Dark regions are the tails where drift is strong, and CTAs make money. The central light region is where CTAs struggle. With higher TAC, we spend less time in the centre.

How valuable (in expectations) to a trend follower is this additional variance?

$$\text{Sharpe}_{3m}(\mu) = \frac{\mu^2/2}{\sqrt{1+\mu^2/4}}$$

The formula behaves somewhere between μ^2 (near zero) and $|\mu|$ (in the tails). Hence the benefit is a *multiplier* of somewhere between 1.1 and 1.2 on its drift-driven expected performance.

It Don't Mean a Thing if it Ain't Got That Swing

Higher autocorrelation of returns helps us as trend-followers in two ways. It makes trends self-reinforcing and increases the likelihood of experiencing stronger underlying trends. Just as importantly, it makes those trends smoother and easier to harvest. This improvement in trend quality raises the expected Sharpe ratio across the board, for both long-term and short-lived trends.

Although it has just been a lot of fun to dance mathematically with you all today, it's time to switch to jazz. We're listening to Charles Mingus right now and thinking, when it comes to trends, like jazz, smooth is the way to go.

Yoav Git
Quant Research

Adam Peddle
Senior Quantitative Researcher

GCM-3485470PE-O0424X

3492072